

## **Assessing the effectiveness of conservation measures: resolving the “wicked” problem of the Steller sea lion**

**I.L. Boyd**

Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews, St Andrews KY16 8LB, UK. [ilb@st-andrews.ac.uk](mailto:ilb@st-andrews.ac.uk)

### **Abstract**

“Wicked” problems are those that are complex and that change when solutions are applied. Many conflicts in conservation fall in to this category. The study approached the problem of how to constrain the apparent wickedness of a problem in the conservation management of a species by using simple empirical indicators to carry out iterative assessment of the risk to a population and to document how this risk evolves in relation to the addition of new data and the implementation of management actions. Effects of high levels of uncertainty within data and also concerning population structure were examined through stochastic simulation and by exploration of scenarios. Historical trends in the example used, the Steller sea lion, showed rapid declines in abundance in some regions during the 1980s. The current total population is 130,000-150,000 Steller sea lions through Alaska and British Columbia and this number has been stable since about 1990 in spite of regional differences in population dynamics. Regional differences in the sequence of changes in the number of pups and non-pups, suggested that an internal re-distribution of juveniles could have happened between 1980 and 1990. Current productivity also appears close to the long term mean. Stochastic population projection using various scenarios showed that, based upon this history, the risk of extinction for the population has declined and is below reasonable thresholds for considering the population to be endangered. The trends in risk suggest that management actions taken since 1990 have probably been effective. Consequently, the conservation management objectives for the Steller sea lion are probably being met. The approach provides a mechanism, based upon experience and scenario analysis, for exploring future policy options and may help to constrain the debate amongst stakeholders about the cost-benefit trade-offs associated with different options.

## Introduction

Conservation management can be costly both in terms of the direct costs of implementation and also because of opportunity costs to social and economic objectives that may be seen to conflict with conservation objectives. Some problems of conservation management can also be described as “wicked” (Jentoft and Chuenpagdee, 2009), in the sense that they are complex, difficult to define and delineate, they have a tendency to reappear and are therefore unconstrained and they tend to change in response to a solution. This is a common feature of problems that lie at the boundary between sociology and biology (Miller, 1993; Goldsmith, 1969). In order to deal with this and to potentially constrain the apparent wickedness of these types of problem, or to “tame” them, it is important to assess the extent to which conservation management is meeting objectives and to proceed as much as possible on the basis of the available information rather than on supposition built upon poorly tested assumptions, or partial information. A challenge for conservation biology is to define clear objectives and to synthesise and simplify the problem of assessing whether conservation objectives are being met. However, in the case of species conservation, the magnitude of this challenge is exacerbated by factors such as the presence of high levels of uncertainty in data, the presence of natural variability in background environmental conditions that could be a major contributor to population trajectories and slow population response as may be the case when considering long-lived species with low intrinsic rates of increase. These types of issues are emerging as challenges to the IUCN Red List criteria used to assess the conservation status of species (Freeman, 2008; Godfrey and Godley, 2008). The present study uses the Steller sea lion (*Eumetopias jubatus*) as an example of such a problem.

The Steller sea lion has been a focus of concern in the North Pacific for at least two decades. Widespread, long-term population declines over important parts of its range have triggered responses from managers in the United States of America under the terms of the Endangered Species Act (ESA). The history of the population decline, population biology and the management response are detailed in the most recent Recovery Plan for the Steller sea lion (National Marine Fisheries Service, 2008). For the purposes of management, the population has been considered to exist as two distinct stocks known as the Western Distinct Population Segment (WDPS) and the Eastern Distinct Population Segment (EDPS) with a dividing line between these stocks at Cape Suckling, Alaska (144°W). The WDPS has been classified as “endangered” and the EDPS has been classified as “threatened” under the ESA. This has placed a duty upon the US National Marine Fisheries Service (NMFS), which is the

responsible agency, to develop and implement a recovery plan for the Steller sea lion.

The presence of the largest commercial fishery in the United States, when measured by volume, mainly for walley pollock (*Theragra chalcogramma*, [http://www.nmfs.noaa.gov/fishwatch/species/walleye\\_pollock.htm](http://www.nmfs.noaa.gov/fishwatch/species/walleye_pollock.htm)), that is broadly spatially coincident with some locations where declines have taken place has added complexity to the management of the Steller sea lion populations. Evidence for an indirect effect of fishing is difficult to find (e.g. National Research Council, 2003; Wolf and Mangel, 2008) and the consensus at present suggests that the population is likely to be tracking a change in the environmental carrying capacity that appears not to be driven by fishing (Trites et al., 2006).

Specific criteria that may need to be addressed when managing wildlife populations are to ensure that the population is maintained at a viable level; that there is an appropriate response to any significant decline; and that the historical range of the species should be maintained. Some criteria may also add the objective of managing populations to be functioning parts of ecosystems. Two problems presented by these criteria, and illustrated by the Steller sea lion, are the underlying assumption of a stationary carrying capacity and the interpretation of historical distribution and abundance as the target level for the species. This has led to the belief that “recovery plans” (e.g. NMFS, 2008) can be created to counteract range contraction or declines in abundance. Some situations, such as deforestation, lend themselves to providing clear evidence of changing carrying capacity but in others, such as the Steller sea lion, this is much less clear (Trites et al., 2006). Because of the way in which conservation legislation is often framed, managers are left with the unenviable task of operating without clear objectives. In the most difficult cases they also have few tools for management other than to speculatively reduce possible conflicts with anthropogenic factors, often setting up costly and socially divisive conflicts. As in the case of the Steller sea lion, where management measures have restricted fisheries activities, the cost-benefit trade-off of conservation measures can be highly uncertain.

An approach often adopted by managers to justify conservation measures is to assess the risk of extinction of a population using population viability analysis (PVA), or similar methods. For long-lived species, a probability of quasi extinction of less than 1% in 100 years has been suggested as the threshold for considering a species to be endangered (Angliss et al., 2002; DeMaster et al., 2004). For the Steller sea lion the quasi-extinction level has been equated to a total population size of 4,743 Steller sea lions (NMFS, 2008).

This number arises if one sets the cut-off using a “genetically effective population size” of 1,000 animals (DeMaster et al., 2004).

PVAs use current knowledge of past population dynamics as well as general life-history information to predict future trajectories. Although PVAs may be broadly satisfactory for predicting extinction risks (Brook et al., 2000), these predictions may be sensitive to factors such as spatial and temporal correlation within the data structures (Melbourne and Hastings, 2008). To date, at least eight models have been developed for Steller sea lions to examine extinction risk (York et al., 1996, 3 models; Gerber and VanBlaricom, 2001, 3 models; Winship and Trites, 2006; NMFS, 2008). All have produced broadly similar predictions. However, the existence of so many PVAs suggests a need for ongoing tracking of extinction risks as a way of assessing the performance of management measures against a consistent standard.

The objectives of the present study were to develop a method for assessing the performance of management against a background of high uncertainty with the ultimate aim of helping focus decision making about management upon robust principles. To achieve this, it was first necessary to collate, assess and analyse the historical population data, a task that had surprisingly not been done for Steller sea lions since the early 1990s (Trites and Larkin, 1996; Loughlin et al., 1992). The analysis focussed upon survey estimates of the number of Steller sea lions at rookeries and haulouts which is by far the richest source of data about the Steller sea lion and for which there are comparable range-wide estimates. This was followed by stochastic simulation of population futures based upon historical data (including its uncertainties) associated with different periods in the management history of the population to assess the risks to the population under different scenarios, to examine how sensitive the risks are to the addition of new data and how well management may have met its objectives.

## **Methods**

### **Count data and definitions**

All data were obtained from the Alaska Fisheries Science Center Steller sea lion database ([http://www.afsc.noaa.gov/nmml/alaska/Steller\\_sea\\_lionhome/databases/](http://www.afsc.noaa.gov/nmml/alaska/Steller_sea_lionhome/databases/)) as well as more recent sources (Fritz et al., 2008; Olesiuk, 2008). Although there are rich data sources for Steller sea lion biology, the data that has been collected most consistently and most reliably are counts of pups at “rookeries” across the species range and counts of non-pups (juveniles and adults) at “haulouts” during the breeding season in June and July. In the present study, no distinction was made between “haulouts”

(sites where animals rest out of the water) or “rookeries” (where pupping takes place) during the breeding season. This is because the historical distinction appears to be somewhat arbitrary, e.g. Pitcher et al. (2007) decided that any site having <50 pups was classified as a “haulout” rather than a “rookery”.

The principal source of data about Steller sea lions came from periodic surveys of these sites. These data have been collected at the resolution of individual sites but had previously been reported at various scales by grouping sites into sub-regions, trend sites, regions and stocks (e.g. York et al., 1996). An important underlying feature of the present analysis is that it was based upon the finest scale of data collection and included all count data, not just the 6 range-wide surveys (e.g. as used in the PVA in NMFS 2008). This meant that the analytical approach was designed to cope with partial coverage in some years and missing values for many sites. Consequently the need for aggregation of sites in to “trend” sites (Fritz et al., 2008) or by region (Sease and Gudmundson, 2002; Pitcher et al., 2007; Fritz et al., 2008; Olesiuk, 2008) was avoided. Any aggregation of sites was based upon clustering derived from congruence between the population dynamics shown by sites (see Supplementary Information for details about the methods used to cluster sites). However, the analysis aggregated sites across the EDPS and WDPS to allow analysis with respect to the historical approaches to management which have recognised these segments as distinct management unit.

Although past studies have fitted population models to current count data (e.g. York et al., 1996; Trites and Winship, 2006; Holmes et al., 2007), the largely unknown and non-stationary error structure within the count data means that the capacity of the data to support these types of analyses is quite limited. Consequently, the present study did not attempt to directly model the population dynamics of Steller sea lions but it did produce an index of productivity called the “pup ratio”, which was the number of pups counted as a proportion of the number of non-pups counted.

### **Observation error**

Various methods have been used to count Steller sea lions. These include, aerial photography using 35mm film format normally using an oblique aspect with hand held cameras and no motion compensation; aerial photography using medium-format (5 inch) with motion compensation; aerial photography using digital format with motion compensation; aerial counts using visual estimates; beach/ground counts; viewpoint observations; counts from a skiff near to shore, and counts from a vessel offshore (Supplementary information,

Table 1). Each method will have different levels of error and in most cases these have not been quantified.

Five general categories of processes were identified that will lead to error in estimates of Steller sea lion abundance. The main processes leading to error in observation are described in detail in the Supplementary information and these include (1) incomplete coverage of the site; (2) variable capacity to observe animals even if they are present; (3) variance due to the counter; (4) dependency on the timing of the survey in relation to both the time of day when it is carried out and the date in relation to the peak of the season; (5) dependency upon the type of terrain. Many of these processes will interact with one-another and all will vary depending upon survey method. The supplementary information evaluates the range of error associated with each of the factors and these ranges were used to generate an estimate of the overall uncertainty around the counts. This was achieved by assuming that all sources of uncertainty were independent and by modifying each count using a uniform distribution across the range of possible values for each source in the Supplementary information (Table 1).

### Modelling future trends

The objective was to predict the range of future population trends conditional on past population trends and current population state. A four-step approach was adopted. Step 1 defined the process by which observations, in this case actual counts of Steller sea lions at each site, represented as a set of state vectors  $\mathbf{x}_k$ , where  $k$  denotes a particular site, were translated in to an equivalent set of state vectors  $\mathbf{n}_k$  representing the real state of the populations at a particular site. As already described, this involved accounting for the errors within the data and building distributions of possible alternatives of  $\mathbf{x}_k$ . It follows from this that

$$N_t = \sum_{k=0}^{k=K} \mathbf{n}_k \quad 1.$$

Where  $N_t$  is the total population size in year  $t$ .

Step 2 defined the potential distribution  $\theta$  of  $\lambda$  which is the proportional change in the state of each successive element in the state vector  $\mathbf{n}_k$ . This defines the historical trends in the population. Assuming exponential growth or decay, the rate of population change ( $\lambda$ ) was derived as follows:

$$n_{k,t} = n_{k,t-\tau} e^{r\tau} \quad 2.$$

Where  $n_{k,t}$  was the number of non-pups or pups in the population at time  $t$ ,  $n_{t-\tau}$  was the number at the previous time of survey, defined by a time interval of  $\tau$

years and  $r$  was the growth rate where  $\lambda = e^r$ . Both  $n_{k,t}$  and  $n_{k,t-\tau}$  were measured variables. The exponential rate of increase was derived as:

$$r = \frac{\ln(n_{k,t}) - \ln(n_{k,t-\tau})}{\tau} + \xi \quad 3.$$

In this case  $\xi$  was a random deviate defined by the standard deviation around the principal eigenvalue of a Lefkovich matrix representing pup, juvenile and adult stages in the population and where the principal eigenvalue was equal to  $\lambda$ . The Weiner process described by Dennis et al. (1991) was used to estimate the value of  $\xi$ . This log-linear relationship, together with uncertainty, was used to interpolate the number of pups and non-pups at each site for years when surveys did not take place.

Step 3 was the process by which the elements within the population state vector at a site,  $n_{k,t}$ , was updated to define scenarios for future population trajectories. This used the distribution  $\theta_n$ , where  $\theta$  was subscripted by  $n$ , because  $\lambda$  varied depending upon the size of the population at each site. Thus the population at each site was

$$n_{k,t} = f(n_{k,t-\tau}, \theta_{n_{k,t}}) \quad 4.$$

In this case,  $\tau$  was equal to unity in the projections. The value of  $\lambda$ , drawn from the distribution  $\theta$  (Equation 4) was selected to satisfy the condition of a first order Markov process in which  $\lambda_{k,t} = \lambda_{k,t-\tau} \alpha_\tau$ , where  $\alpha$  was a random variate drawn from a distribution defined by the direct observation of the autocorrelation between values of  $\lambda$  at different values of  $\tau$ .

Step 4 involved aggregating the results for sites across different scales of relevance to management and assessing the probability of recovery criteria being met.

### Density limitation

There was no direct information available about the carrying capacity. However, it can be assumed that there are upper limits to the number of Steller sea lions at particular sites, because of limits on space or because of limits on the local environmental productivity. To allow for this type of limit within the model, it was assumed that each site had an upper limit of population size drawn from the distribution of maximum observed values across all sites. Since the population at many sites have only been measured while they have been increasing, this approach will tend to produce a conservative estimate of overall carrying capacity.

### Model definitions

Several models were used to investigate possible future trends in the population and these are summarised in Table 1. Spatial correlation of sites was included in Model 2 but not in Model 1. Both models were run on data for the whole population and also separately for the Western and Eastern distinct population segments (WDPS and EDPS respectively) (York et al., 1996; NMFS, 2008). Results from all these combinations of models were investigated using data from the complete time series (a); from 1990 to the present (b) and from 2000 to the present (c). These data sets in (b) and (c) represent two different phases of conservation management. Before 1990 there were no conservation management measures; after 1990 measures to reduce human impacts were introduced and after 2000 measures to reduce possible indirect effects of fishing were expanded. Consequently, population projection using data from 1990 and 2000 respectively assumed similar types of management will be extended in to the future.

[Table 1 here]

## **Results**

### **Data available**

The complete data set contained a sample of 7,175 counts of non-pups and 906 counts of pups at individual sites. After exclusion of data outside the June/July pupping season and taking averages for repeated counting at individual sites within years, the total sample of site counts for non-pups was 4,530 and 828 for pups. The total sample of sites from the region included within the present analysis for non-pups was 431 and the total number of rookeries was 312.

### **Observation errors**

Taking information from the whole data set where there were repeated counts of a site within a single year, the coefficient of variation (CV) followed a log-normal distribution with a mean of 0.28 (sd = 2.492). The distribution of CVs (Fig. 1) is unlikely to have completely captured the biases associated with the characteristics of the different sites because it tested for internal consistency in within-site counts and says little about consistency among sites. Unfortunately, there was very little that could be done to assess the differences in count bias among sites even though it is possible that this could be important. For example, Cumbley et al. (1997) showed clear differences in the seasonal pattern of abundance among neighbouring beaches on Marmot Island. There was also evidence of large inter-site variation in the apparent



proportion of non-pups hauled out at sites during the breeding season. The local conditions that drive these differences are not understood but all may add significantly to the overall uncertainty around counts.

**[Figure 1 here]**

The overall effect of the procedure used to account for error and bias is shown in Figure 2 for non-pup counts. This shows how the true population is likely to be distributed as a multiple of observed population.

**[Figure 2 here]**

### **Historical abundance**

The original interpolated counts across the whole population, together with counts corrected for error and bias are shown in Figures 3a & 3d. Although the number of pup and non-pup Steller sea lions counted apparently increased between 1960 and 1980, relatively few sites were surveyed through this period so much of this increase could be caused by increased observation effort. The total number of pups and non-pups declined after 1980 even though the observation effort, in terms of number of sites included within the surveys, continued to increase.

The number of non-pups counted declined almost linearly through the 1980s so that, by 1990 <50% of the numbers observed at the peak in 1980 were being counted (Figure 3a). However, from 1990, the total number of non-pups has remained stable and the apparent observation effort based on the number of sites included in the counts has also changed relatively little through this time.

The number of pups observed did not begin a sustained decline until after 1985, and suggested a roughly 5-year delay between declines in non-pup abundance and declines in pup abundance (Figure 3d). This decline continued until about 2000 and thereafter numbers counted have increased at about the same rate as the previous decline. The rapid increase in the number of sites surveyed for pups since 2000 partly reflects greater observation effort but it also reflects a reclassification of some sites into a larger number of smaller units.

The application of the procedure to account for error in the counting methods caused an increase in the estimated number of pups and non-pups. As expected from the distribution illustrated in Figure 2, accounting for counting errors led to an approximate doubling of the number of non-pups (Figure 3a).

The confidence limits around the mean estimate narrowed through time reflecting the increasing accuracy of the counting methods used and the relative level of uncertainty was greater for pups than non-pups. Thus, based on this account, the total number of non-pup Steller sea lions within Alaska and British Columbia is between 130,000 and 150,000. The trends in these estimates follow those for the observed numbers.

The “pup ratio” before 1980 was excluded as unreliable (Fig. 3g). These data suggest that relative productivity increased through the 1980s, when the population was in decline. The pup ratio then declined through the 1990s when the population had stabilised and began to increase again after 2000.

The Eastern and Western Distinct Population Segments (Figs. 3b, 3c, 3e & 3f) followed roughly opposite trends (Fig. 4). The rapid decline in abundance through the 1980s in the WDPS meant that the population reached about 20% of its peak size and it has shown a much slower overall decline since then (Fig. 3c). The decline in pup numbers in the western segment of the population lagged behind the change in the population as a whole (Fig. 3f) and this is reflected in the increasing pup ratio through the 1980s. The pup ratio in the western population segment then declined rapidly between about 1990 and 1997 and has remained unchanged at an intermediate level, or has shown a slight overall increase (Fig. 3i).

In contrast to the western segment, the number of non-pups and pups in the eastern segment has shown a near-monotonic increase since before 1980 (Figs. 3b & 3e). The pup ratio in the eastern segment declined through the 1980s (Fig. 3h), at a time when the pup ratio in the west was high or increasing and when the non-pup portion of the population was in rapid decline. The pup ratio in the east has remained relatively low or increased slightly since 1990, but at a level equivalent to the pup ratio in the western population segment in recent years. The apparent low pup ratios in the eastern population segment during the early 2000s could have been caused by unaccounted bias within a single survey.

**[Figure 3 here]**

**[Figure 4 here]**

### **Rate of increase ( $\lambda$ ) and the distribution of $\lambda$**

The overall rate of increase at sites was heavily centred on unity (Fig. 5) which reflects the fact that most of the data came from a population that has been showing little directional change in recent years. The distribution was highly

leptokurtic. Although this type of distribution could potentially be modelled using the Parieto or Generalised Extreme Value distributions, to avoid additional uncertainty due to model fitting, bootstrap re-sampling from the empirical distribution was used to model future population growth at each site.

**[Figure 5 here]**

However, the rate of increase appeared to be affected by population size in that increasing population size resulted in lower rates of increase (Kruskal-Wallis, Chi-square= 46.07,  $df=10$ ,  $p<0.001$ ). Population size had a particularly strong effect upon the variance in  $\lambda$ . This was mainly because of high variability in  $\lambda$  at low population sizes presumably because of an increasing effect of immigration and emigration and stochasticity in the data. Consequently, population size was used as a covariate in the choice of rates of increase to model future population growth (see Equation 4).

There was strong evidence that the rate of increase at individual sites was dependent upon the rate of increase at the previous time of measurement (Fig. 6). This autocorrelation was also included in population projections.

**[Figure 6 here]**

### **Future trends in abundance**

#### **Model 1**

The results of population projections using Model 1, which included temporal correlation but no spatial correlation (Fig. 7), showed the risk of extinction depended upon which parts of the time-series of historical counts was used to model the rate of change in the population. The population as a whole (Fig. 7a) when modelled using all historical data had a probability of extinction that exceeded the management objective until 1993 and after that the probability of extinction was  $<0.1$  in 100 years (referred to here as the “conservation objective”). The terminal population size after 100 years was considerably greater in the case where no density limits were used and this pattern was the same across the population segments and irrespective of which prior distribution of rates was used (contrasting Models a, b and c). The western population segment (WDPS, Fig 7e) showed less of a tendency to meet the conservation objectives early in the time series but using the full range of historical data, it was compliant with the conservation objective after 2005. Using Model 1 and only the later parts of the time series of  $\lambda$  to project the population forward showed that in no circumstances did the stochasticity

included here result in the extinction of the populations (Models 1 and 1c; Figs. 7b, 7d and 7f). The terminal size of the population as a whole (Fig. 7a) and of the western stock (Fig. 7c) continued to increase as additional data were added, reflecting the addition of information about the population as it has stabilised through time (Figs. 3a and 3c) and suggesting that this period of stability is increasing confidence that the population is robust to extinction. This analysis suggested that the predicted risk to the population will be relatively insensitive to the addition of new data.

## Model 2

The results of population projections using Model 2, which included temporal correlation and spatial correlation, are shown in Figure 8. Overall, the addition of correlation increased the probability that the population would not meet the conservation objective but this effect was only present in Model 2a (Fig. 8a, 8c & 8d) and not in Models 2b and 2c, and the present probability of not meeting the conservation objective was very low. Adding density limits produced an increased probability that the populations would not meet the conservation objective and this is illustrated by the pattern in Figure 8c where even the eastern population segment failed to meet the conservation objective up to 1995. As with Model 1, the total terminal populations after 100 years were predicted to increase to a roughly constant level, showing that the prediction appears to be robust to the duration of the time series used to derive the sample distributions of  $\lambda$ .

Use of the time series from 1990 (Phase A in the management history, Figs 8b, 8d and 8f) showed slightly elevated terminal populations compared with the scenarios using the complete time series. The terminal population levels derived using only the data from Phase B of the management history (post 2000) were similar to those derived from the distribution of  $\lambda$  from Phase A of the management history. When the population changes during Phases A and B were used, all scenarios suggested that the population met the conservation objective.

**[Figure 7 here]**

**[Figure 8 here]**

## Discussion

The Steller sea lion is a particularly challenging example for testing methods to assess conservation status. However, the present analysis has demonstrated using a broad range of different scenarios and reflecting different

combinations of assumptions, including reasonable estimates of uncertainty in data, that it is possible to derive a coherent assessment that may release managers from the yoke of over-precaution (Gillespie 2007).

### **Observation process**

Count data contained errors only some of which have been taken in to account in past analyses of population trends (e.g. NMFS, 2008; Fritz et al., 2008). Past analyses have either accepted count data without error or with generalised error terms used when fitting models (e.g. Winship and Trites, 2006; Wolf and Mangel, 2008). The error defined here was approximately log-normal (Fig. 2) but the mean and variance of this distribution will vary with the counting methods used and has changed through time. Overall, the present analysis suggests that past approaches to understanding the population dynamics of the Steller sea lion have generally underestimated the uncertainties within the raw population data.

There has been an accumulation in the number of sites surveyed through time. We cannot be sure what proportion of this accumulation reflects the dispersal process as sea lions move to occupy new rookeries and haulouts and what proportion represents increasing diligence on the part of those conducting surveys. However, it is unlikely that the relative number of sites surveyed between the start and end of the time series (Fig. 3) is directly proportional to the observation effort.

### **Data considerations**

Time-series count data from marine mammals are frequently complex and difficult to interpret mainly because it is often difficult to carry out truly synoptic surveys of widely spread marine mammal populations, not all animals are available to be surveyed, the relationship between the observed number and the true number is often obscure and difficult to assess, and there are often missing values caused by stochastic factors such as weather, funding, equipment failure, platform availability and, in the case of the Steller sea lion, withdrawal of permits to conduct surveys. More often than not, those involved in surveys have to make a leap of faith that the numbers they count are at least internally consistent because quantitative assessment of all the errors that could occur is extremely difficult. However, the type of post-hoc assessment of the scale of error applied here can provide additional information that is relevant to management.

The approach adopted in the present study using sites as the unit of sampling overcomes having to manage incomplete surveys or surveys that have a

regional focus that shifted between years, often by excluding the data from the assessment of the performance of conservation measures (e.g. NMFS, 2008). However, even in the present example the data set did not include sites from the whole population: data from Washington, Oregon or those from Russia were not included within the NMFS data base. It would be simple to include these in future if data were made available at the level of individual sites. Lack of these data is unlikely to have seriously affected the outcome of this analysis because these regions make a comparatively small contribution to the overall population (Pitcher et al., 2006; NMFS, 2008). Future synoptic analyses of the Steller sea lion population would be simplified if all data holders could contribute to a common data base.

The analysis also allowed the data to dictate the cluster structure rather than geography, genetics or past perceptions of population sub-divisions. A number of analyses have sought to sub-divide the population demographically, geographical, politically and using genetics. O'Corry-Crowe et al., (2006) showed high levels of mitochondrial DNA diversity within Steller sea lions but, because of the recent trajectory of the population, it is not possible to know whether the distribution of haplotypes in the population is a cause or a consequence of these demographic changes. Consequently, the present study has not made any underlying assumptions about the structure of the population. However, because of the historical management approach taken to Steller sea lions, the effects of considering the eastern and western segments separately has been investigated, but there seems to be no *a priori* reason, other than historical precedence, to view these population segments as distinct.

### **Historical trends in abundance**

The data presented in Figure 3 need to be interpreted with care in terms of fine-scale changes shown. Even the estimates adjusted for error cannot account for the relatively low observer effort in the early part of the time series. It is likely that information from before 1980 contains an underestimate of the total population, although we do not know this for certain. The results of the present assessment of the Steller sea lion population trends before 1980 are in broad agreement with earlier estimates (Loughlin et al., 1984; Loughlin et al., 1992; Trites and Larkin 1992). The present analysis indicates that Steller sea lion numbers may have increased through the 1960s and 1970s. It appears that some status reviews of Steller sea lions (e.g. NMFS, 2008) and some PVAs (e.g. York et al., 1996; Gerber and VanBlaricom, 2001; NMFS, 2008) reflect an assumption that the counts of the population before 1980 greatly underestimated the overall population and these PVAs appear to have used the counts after 1980 as a benchmark for the pre-1980 estimates of Steller sea

lion numbers, even though there is little evidence for this within the available data. After 1980, the non-pup population declined consistently for 10 years from a high of ~312,000 in 1980 to a low of ~135,000 in 1990, an annual rate of decline of 8%. There is no evidence of a significant decline or increase in the overall non-pup population since this time. However, local or regional dynamics may differ (e.g. York, 1994; Trites and Larkin, 1996) meaning that it is difficult to draw conclusions about population status based upon regionally-based surveys or assessments focussed upon sub-regions.

There are potentially important implications of these historical data for setting management objectives. It appears that the population size in 1980 could have been a peak and could have been greater than the long-term carrying capacity. Consequently, the return of the population since 1990 to a level similar to that measured in the 1960s may represent a level that is closer to the long-term mean carrying capacity. This suggests that management objectives could be most usefully focused upon maintaining the current population level rather than attempting to “recover” the population towards an historic high (NMFS, 2008), unless it can be shown that the environment has returned to the conditions of the 1970s (see Trites et al., 2006).

Overall, the population declined in the west and increased in the eastern part of its range (Fig. 3). While data from genetics, mark-recapture, and satellite tagging (Raum-Suryan et al., 2002; O’Corry-Crowe et al., 2006; NMFS, 2008; Lander et al., 2009) do not suggest that there are regular movements of individuals from west to east, it remains likely that these opposing trends reflect a re-distribution of favoured habitat and a shift in the distribution of food. The pup ratios (Fig. 3) may indicate changing population productivity but they could also indicate movement of non-pups, most likely juveniles. It may be significant that the pup ratios in the eastern segment of the population declined at the same time as they increased in the western segment and at a time when the absolute size of the non-pup population in the eastern segment of the population began to increase (Fig. 3h compared with Fig. 3i and Fig. 3f). Rather than interpreting these as changes caused by internal dynamics within the eastern and western segments of the population, these could indicate recent emigration of juveniles from the western segment with subsequent recruitment in to the eastern segment. This hypothesis could be tested by reassessing the genetics data against the hypothesis generated by the present study, although since few samples are likely to be available that pre-date the changes in the 1980s even this may be inconclusive. The outcome has implications for the approaches to management of the Steller sea lion as two separate stocks because of the effects that dispersal could have on meeting management objectives (Taylor, 1997).

Some features of the population data may reflect the potential causes the decline in the Steller sea lion population during the 1980s. The lag of 5 years between the start of the decline in pups compared to adults and the increase in the pup ratio through the same period supports the hypothesis (York, 1993; Holmes and York, 2003) that the decline through the 1980s was largely the consequence of low juvenile survival, but it could also have been caused by the emigration of juveniles (see above). This trend appears to have been driven by changes in the west of the range. In other words, the time lag in the decline in pup production was probably a consequence of the loss of females recruiting to the adult population and the increasing pup ratio in the early 1980s was probably because adults remained relatively unaffected. However, the ratio of pups in the population then declined and this suggests that adult fecundity declined through the 1990s, possibly as a delayed effect of the processes going on in the 1980s. Nevertheless, the subsequent increase in the pup ratio, with the apparent increase in the non-pup population, suggests that fecundity is once again recovering. This conclusion contrasts with that of Holmes et al., (2007) who suggest that fecundity is lower than expected and probably warrants further investigation. Differences may be caused by the smaller spatial scales used by Holmes et al., (2007) and we know that the population dynamics can be volatile at these scales (York, 1994; Trites and Larkin, 1996).

### **Scenarios for future trends**

Simulations of future trends in abundance showed that if one assumes that there is a stationary distribution of  $\lambda$  across the full range of the data available, the probability of the population as a whole meeting the conservation objectives (probability of extinction  $<0.01$  in 100 years) was met under all circumstances (Figs 7 & 8). The density-limited cases of both Model 1 and Model 2 constrained the growth of the population within reasonable boundaries and, even in these circumstances, the population met the conservation objectives. Only when distributions of  $\lambda$  that excluded observations for the past 10 years were used were the conservation objective not met. Model 2 included a more complex but potentially more realistic representation of the population processes. Inclusion of information about metapopulation structure, as opposed to considering all sites as independent populations, tended to increase the vulnerability of the population to extinction. This result is similar to that found by Melbourne and Hastings (2008). Apparent vulnerability to extinction will increase further if the Eastern and Western segments of the population are considered to be internally freely-mixing but separate. This has often been assumed within past studies that have modelled the extinction probability of the Steller sea lion in terms of eastern and western distinct population segments (e.g. Gerber and



VanBlaricom, 2001; NMFS, 2008), even though genetic studies have tended to indicate that these sub-sections of the population are not freely mixing (Bickham et al., 1998; O'Corry-Crowe et al., 2006; Hoffman et al., 2006) and, as indicated by the present study, they may not be separate. Consequently, making the assumption that the eastern and western segments are freely mixing sub-populations is likely to result in an unrealistically inflated estimate of the extinction probability calling in to question the results of some past risk assessments of the Steller sea lion population.

Although there was some level of density dependence included in the model, because the mean value of  $\lambda$  declined as population size increased, the only explicit form of density regulation in these models was the upper limit set on the number of sea lions that could be present on each site. This was introduced as an additional stochastic variable so the predictions will have accounted for uncertainty in the effects of density. There is no information from Steller sea lions about how density-dependent regulation actually operates but the overall effect of limiting the total number of sea lions within a site is likely to be a reasonable surrogate.

There are reasons to believe that the distribution of  $\lambda$  was not stationary. Inspection of the population trends in Figure 3 suggest that a different set of conditions applied to the population before and after 1980 and before and after 1990. The threshold at 1990 is further supported by the introduction of extensive conservation measures at that time, and further measures were introduced after 2000 (NMFS, 2008). Consequently, if it is assumed that management approaches similar to the current conservation measures are maintained, then repeating the population projections using the distribution of  $\lambda$  from after 1990 and after 2000 are likely to provide a more realistic prediction of future risks. This demonstrated that the Steller sea lion population as a whole, or if considered in the two segments, met the conservation objectives (Fig. 7b, 7d, 7f, 8b, 8d and 8f).

### **Implications for conservation actions**

The conclusions of previous assessments of the risks to the Steller sea lion population were based upon the data that were available at the time they were developed (e.g. York et al., 1996). Although none of the previous studies has examined the population as a whole (preferring instead to consider it as two distinct populations or even sub-sections of these populations), the results of the present study suggest that both the eastern and western segments of the population have probabilities of persistence that mean they do not meet the criteria for classification as endangered and it would be reasonable to de-list them. They also suggest that conservation actions

undertaken to date have either been successful or neutral in their effect. Even if one takes the most precautionary approach by assuming that management actions have had no positive effect the best risk models to apply would be Models 1a and 2a that have density limits in place. Both of these demonstrate compliance with conservation objectives.

### **General conclusions**

Two general messages for conservation can be derived from the specific example of the Steller sea lion. First, the present study has demonstrated that the success of management can be assessed by updating analyses of risk with new data within a stochastic framework, using scenarios built upon the empirical distributions of directly measured population indicators. This contrasts with approaches using complex process models of the population or by the making assessments based only upon the latest information, which has often been the case for the Steller sea lion. The integration of historical information into a framework that makes few unsupported assumptions and that examines any assumptions through the stochastic structure of the approach and by the exploration of scenarios is open to modification and update in a manner that can engage with stakeholders and address, and test, their concerns. Second, considering that there will be cost-benefit trade-offs associated with conservation measures, the conservation benefits accruing from additional management actions could be negligible. For example, in the case of the Steller sea lion continuation of current management would appear to be justified but the addition of new management bringing additional social, economic or implementation costs would appear not to be justified. Without ongoing assessment of the likelihood of management meeting conservation objectives – which has not happened with the Steller sea lion as new data have been obtained - it is difficult to make an appropriate assessment of the cost-benefit trade-offs of future management policy. Lack of this type of assessment means that uncertainty in decision-making is not reduced as much as it could be and this encourages unnecessary conflict amongst stakeholders about future policy and it does not “tame” the problem. Consequently, the approach examined in the present study has the potential to both reduce conflict amongst stakeholders and to provide a mechanism for exploring future policy options based upon experience.

### **Acknowledgements**

I wish to thank staff at the US National Marine Mammal Laboratory, Seattle for their assistance and discussions concerning the dynamics of Steller sea lion populations. In particular, I thank Lowell Fritz, Tom Loughlin, Ann York for

discussions over many years. I am also grateful to Andrew Trites from the University of British Columbia for similar discussions and to Peter Olesiuk of the Department of Fisheries and Oceans, Canada for sending data. This study has been funded by many sources over the years including assistance from the US National Marine Fisheries Service, the Marine Conservation Alliance and the UK Natural Environment Research Council.

## References

- Angliss, R.P., Silber, G.R., Merrick, R. 2002. Report of a workshop on developing recovery criteria for large whale species. NOAA Technical Memorandum NMFSF/OPR-21.
- Bickham, J. W., Loughlin, T.R., Calkins, D. G., Wickliffe, J. K., Patton J.C., 1998. Genetic variability and population decline in Steller sea lions from the Gulf of Alaska. *J. Mammal.* 79, 1390-1395.
- Brook B.W., O'Grady J.J., Chapman A.P., Burgman M.A., Akcakaya H.R., Frankham R., 2000. Predictive accuracy of population viability analysis in conservation biology. *Nature* 404, 385-387.
- Chumbley, K.J., Sease, J., Strick, M., Towell, R., 1997. Field studies of Steller sea lions (*Eumetopias jubatus*) at Marmot Island, Alaska, 1979 through 1994. U.S. Dep. Commer., NOAA Tech, Memo. NMFS-AFSC-77, 99pp.
- DeMaster, D., Angliss, R., Cochrane, J., Mace, P., Merrick, R., Miller, M., Rumsey, S., Taylor, B., Thompson, G., Waples, R., 2004. Recommendations to NOAA Fisheries: ESA Listing Criteria by the Quantitative Working Group. *NOAA Technical Memorandum NMFS-F/SPO-67*.
- Dennis, B., Minholland, P.L., Scott, J.M., 1991. Estimation of growth and extinction parameters for endangered species. *Ecol. Monogr.* 61: 115-143.
- Freeman, M.M.R., 2008. Challenges of assessing cetacean population recovery and conservation status. *Endang. Species Res.* 6, 173-184.
- Fritz, L., Lynn, M., Kunisch, E., Sweeney, K. 2008. Aerial, ship and land-based surveys of Steller sea lions (*Eumetopias jubatus*) in Alaska, June and July 2005-2007. *NOAA Technical Memorandum NMFS-AFSC-183*.
- Gillespie, A. 2007. The precautionary principle in the Twenty-First Century: a case study of noise pollution in the ocean. *Int. J. Mar. Coast. Law* 22, 61-87.
- Gerber, L. R., VanBlaricom, G.R., 2001. Implications of three viability models for the conservation status of the western population of Steller sea lions (*Eumetopias jubatus*). *Biol. Conserv.* 102, 261-269.
- Godfrey, M.H. and Godley, B.J., 2008. Seeing past the red: flawed IUCN global listings for sea turtles. *Endang. Species Res.* 6, 155-159.

Goldsmith, J.R., 1969. Air pollution epidemiology. A wicked problem, an informational maze, and a professional responsibility. *Arch. Environ. Health* 18, 516-522

Hennen, D.R. 2006 Associations between the Alaska Steller sea lion decline and commercial fisheries. *Ecol. Appl.* 16, 704-717.

Hoffman, J. I., Matson, C.W., Amos, W., Loughlin, T.R., Bickham, J.W., 2006. Deep genetic subdivision within a continuously distributed and highly vagile marine mammal, the Steller's sea lion (*Eumetopias jubatus*). *Mol. Ecol.* 15, 2821-2832

Holmes, E.E., Fritz, L.W., York, A.E., Sweeney, K. 2007. Age-structured modelling reveals long-term declines in natality of western Steller sea lions. *Ecol. Appl.* 17, 2214-2232.

Holmes, E.E., York, A.E. 2003. Using age structure to detect impacts on threatened populations: a case study with Steller sea lions. *Cons. Biol.* 17, 1794-1806.

Jentoft, S. and Chuenpagdee, R. 2009. Fisheries and coastal governance as a wicked problem. *Marine Policy* 33, 553-560.

Lander, M. E., Loughlin, T.R., Logsdon, M.G., VanBlaricom, G.R., Fadley, B.S., Fritz, L., 2009. Regional differences in the spatial and temporal heterogeneity of oceanographic habitat used by Steller sea lions. *Ecol. Appl.* 19, 1645-1659.

Loughlin, T. R., Perlov, A.S., Vladimirov, V.A., 1992. Range-wide survey and estimation of total number of Steller sea lions in 1989. *Mar. Mamm. Sci.* 8, 220-239.

Loughlin, T. R., Rugh, D.J., Fiscus, C.H. 1984. Northern sea lion distribution and abundance: 1956-1980. *J. Wildl. Manage.* 48, 729-740.

Loughlin T.R., Perlov A.S., Vladimirov V.A., 1992. Rangewide survey and estimation of total number of Steller sea lions in 1989. *Mar. Mamm. Sci.* 8, 220-239.

Melbourne, B.A., Hastings, A., 2008. Extinction risk depends strongly on factors contributing to stochasticity. *Nature* 454, 100-103.

Miller, A., 1993. The role of analytical science in natural-resource decision-making. *Environ. Manage.* 17, 563-574.

National Marine Fisheries Service. 2008. Recover plan for the Steller sea lion (*Eumetopias jubatus*). Revision. National Marine Fisheries Service, Silver Spring, MD. 325 pp.

National Research Council 2003. The decline of the Steller sea lion in Alaskan waters: untangling food webs and fishing nets. National Research Council, Washington , D.C.. 216 pp.

O'Corry-Crowe, G., Taylor, B.L., Gelatt, T., Loughlin, T.R., Bickham, J., Basterretche, M., Pitcher, K.W., DeMaster, D.P., 2006. Demographic independence along ecosystem boundaries in Steller sea lions revealed by mtDNA analysis: implications for management of an endangered species. Can. J. Zool. 84, 1796-1809.

Olesiuk, P.F., 2008. Abundance of Steller sea lions (*Eumetopias jubatus*) in British Columbia. Canadian Science Advisory Secretariat, Research Document 2008/063, 29pp.

Olesiuk, P.F., Calkin, D.G., Pitcher, K.W., Perryman, W.L., Stinchcomb, C., Lynn, M., 2008. An evaluation of Steller sea lion (*Eumetopias jubatus*) pup counts from 35mm oblique images. Canadian Science Advisory Secretariat, Research Document 2008/064, 42pp.

Pascual, M.A, Adkison, M.D., 1994. The Decline of the Steller Sea Lion in the Northeast Pacific: Demography, Harvest or Environment? Ecol. Appl. 4, 393-403.

Pelletier, F., Clutton-Brock, T., Pemberton, J., Tuljapulkpar, S. & Coulson, T., 2007. The evolutionary demography of ecological change: linking trait variation and population growth. Science 315, 1571-1574.

Pitcher, K.W., Burkanov, D.G., Calkins, D.G., LeBoeuf, B.J., Mamaev, E.G., Merrick, R.L. & Pendleton G.W., 2001. Spatial and temporal variation in the timing of births of Steller sea lions. J. Mammal. 82, 1047-1053.

Pitcher, K.W., Olesiuk, P.F., Brown, R.F., Lowry, M.S., Jeffries, S.J., Sease, J.L., Perryman, W.L., Stinchcomb, C.E. & Lowry, L.F. 2006. Abundance and distribution of eastern North Pacific Steller sea lion (*Eumetopias jubatus*) population. Fish. Bull. 105, 102-115.

Raum-Suryan, K., Pitcher, K.W., Calkins, D.G., Sease, J.L. & Loughlin, T.R., 2002. Dispersal, rookery fidelity, and metapopulation structure of Steller sea lions (*Eumetopias jubatus*) in an increasing and decreasing population in Alaska. Mar. Mamm. Sci. 18, 746-764.

Sease, J.L., Gudmundson, C.J. 2002. Aerial and land-based surveys of Steller sea lions (*Eumetopias jubatus*) from the western stock in Alaska, June and July 2001 and 2002. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-131, 145pp.

Taylor, B.L 1997. Defining "population" to meet management objectives for marine mammals. Mar. Mamm. Sci. 3, 49-65.

Trites, A. W., Larkin, P.A. 1996. Changes in the abundance of Steller sea lions (*Eumetopias jubatus*) in Alaska from 1956 to 1992: how many were there? Aquat. Mamm. 22, 153-166.

Trites, A. W., Miller, A.J., Maschner, H.D.G., Alexander, M.A., Bograd, S.J., Calder, J.A., Capotondi, A., Coyle, K.O., Lorenzo, E.D., Finney, B.P., Gregr, E.J., Grosch, C.E., Hare, S.R., Hunt, G.L., Jahncke, J., Kachel, N.B., Kim, H.-J., Ladd, C., Mantua, N.J., Marzban, C., Maslowski, W., Mendelssohn, R., Neilson, D.J., Okkonen, S.R., Overland, J.E., Reedy-Maschner, K.L., Royer, T.C. Schwing, F.B., Wang, J.X.L., Winship, A.J., 2006. Bottom-up forcing and the decline of Steller sea lions in Alaska: assessing the ocean climate hypothesis. Fish. Oceanog. 16: 46-67.

Winship, A.J., Trites, A.W., 2006. Risk of extirpation of Steller sea lions in the Gulf of Alaska and Aleutian Islands: a population viability analysis based on alternative hypotheses for why sea lions declined in western Alaska. Mar. Mamm. Sci. 22, 124-155.

Wolf, N., Mangel, M., 2008. Multiple hypothesis testing and the declining-population paradigm in Steller sea lions. Ecol. Appl. 18, 1932-1955.

York, A. 1994., The population dynamics of the northern sea lions, 1975-85. Mar. Mamma. Sci. 10, 38-51.

York, A. E., Merrick, R.L., Loughlin, T.R., 1996. An analysis of the Steller sea lion metapopulation in Alaska, in: D. R. McCullough (ed), Metapopulations and wildlife conservation. Island Press, Washington DC and Covelo, CA, pp 259-292.

## Figure Legends

Figure 1. The frequency distribution (histogram) of the coefficient of variation found for repeated counts of non-pups at the same site in the same breeding season. The skewed form of the distribution meant that counting error was modelled using a log-normal distribution (solid line plotted as a proportional frequency scaled to the observed frequency). The geometric mean was 0.280 (sd = 2.492).

Figure 2. The frequency distribution of  $N_{true}$  for non-pup counts expressed as a proportion of  $N_{obs}$  for non-pup counts.

Figure 3. Changes in the observed number of non-pup (a, b, c) and pup (d, e, f) Steller sea lions at rookeries and haulouts during the June-July pupping season (shown as a broad grey line to reflect that there is uncertainty about these numbers). Log-linear interpolation with Weiner diffusion was used to account for missing values. The dots show equivalent numbers adjusted for errors in the observation process. The solid lines above and below these dots show  $\pm 95\%$  confidence intervals around these mean values. The dashed line shows the number of sites included in the estimation during each year. In panels g, h, and i the ratio of pups in the population to non-pups is shown  $\pm 1$  SD. The shaded blocks show the two phases of conservation management of Steller sea lions.

Figure 4. The relationship between the total number of Steller sea lions estimated within in the western and eastern parts of their range. Data from 1980-2006.

Figure 5. The frequency distribution of  $\lambda$  for non-pups at rookery sites during June and July across all sites irrespective of population size (n=2874).

Figure 6. The relationship between the rate of increase at each site and the change in the rate of increase the next time it was measured. Extreme values at each end of these distributions are attributable to fluctuations at small population size, which was a significant co-variate in this relationship.  $\text{Ln Rate difference} = -0.0729 - 1.2138 \ln(\lambda)$ ;  $r^2=0.642$ .

Figure 7. The evolution of population projections for 100 years in to the future using Model 1a (see Table 1 for the definition of the model and results are in panels a, c and e), and Models 1b and 1c (panels b, d and f). The projections used historical time-series that included all the available historical data up to the year in which the time series was terminated (shown on the x-axis) to



illustrate how the predictions evolved as new data were collected. The probability of extinction during the future 100 years is shown as dots and the terminal population size is shown as a mean (white lines)  $\pm$  1 standard deviation (shaded regions). The result of running the models without density limits at individual sites (light grey shading, grey dots) is contrasted with running the models using density limits (dark grey shading, black dots). The projections using Model 1b used historical time-series that included data collected after 1990 when management measures were introduced to promote conservation of the Steller sea lion (phases A & B in Fig. 4). The projection using Model 1c used historical time-series from 2000 to 2006 (phase B in Fig. 4) and the output is shown as a single shaded square ( $\pm$  1 standard deviation) for the predicted terminal population size and an unshaded square to show the probability of extinction. The output for Model 1c with density limits in place is not shown. EDPS signifies the Eastern stock of Steller sea lions and WDPS signifies the western stock.

Figure 8. The evolution of population projections for 100 years in to the future using Model 2a (panels a, c and e), and Models 2b and 2c (panels b, d and f). Refer to the caption for Figure 7 for further details about the meaning of each symbol.

	Model 1			Model 2		
	a	b	c	a	b	c
Spatial clustering included					✓	
Temporal correlation included		✓			✓	
Basic data at the scale of rookeries and haulouts (June/July)		✓			✓	
Sampled $\lambda$ from empirical distribution across the all years	✓			✓		
Sampled $\lambda$ from empirical distribution across phase A and B only (post 1990)		✓			✓	
Sampled $\lambda$ from empirical distribution across B only (post 2000)			✓			✓
Results modified to reflect the observation process, including error and bias		✓			✓	

Table 1. Definitions of the different population risk analyses undertaken. All of these scenarios were investigated with and without density limits.

Figure 1

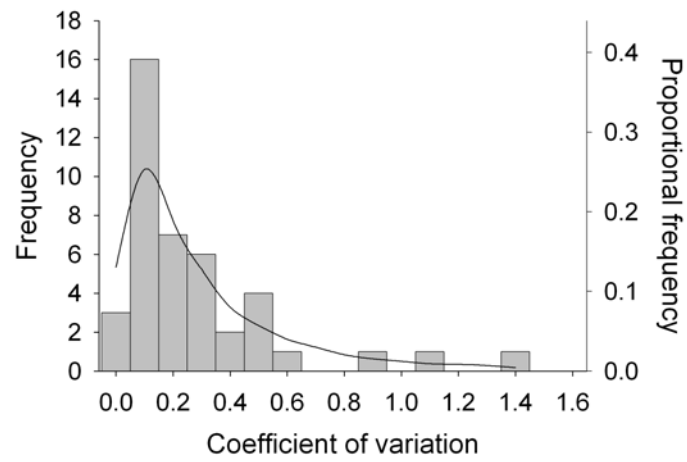


Figure 2

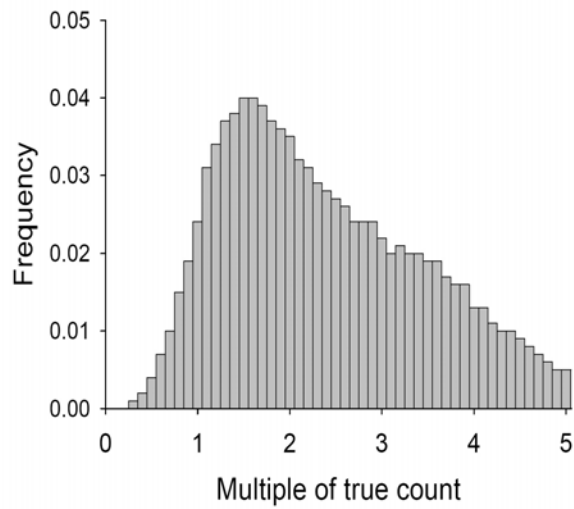


Figure 3

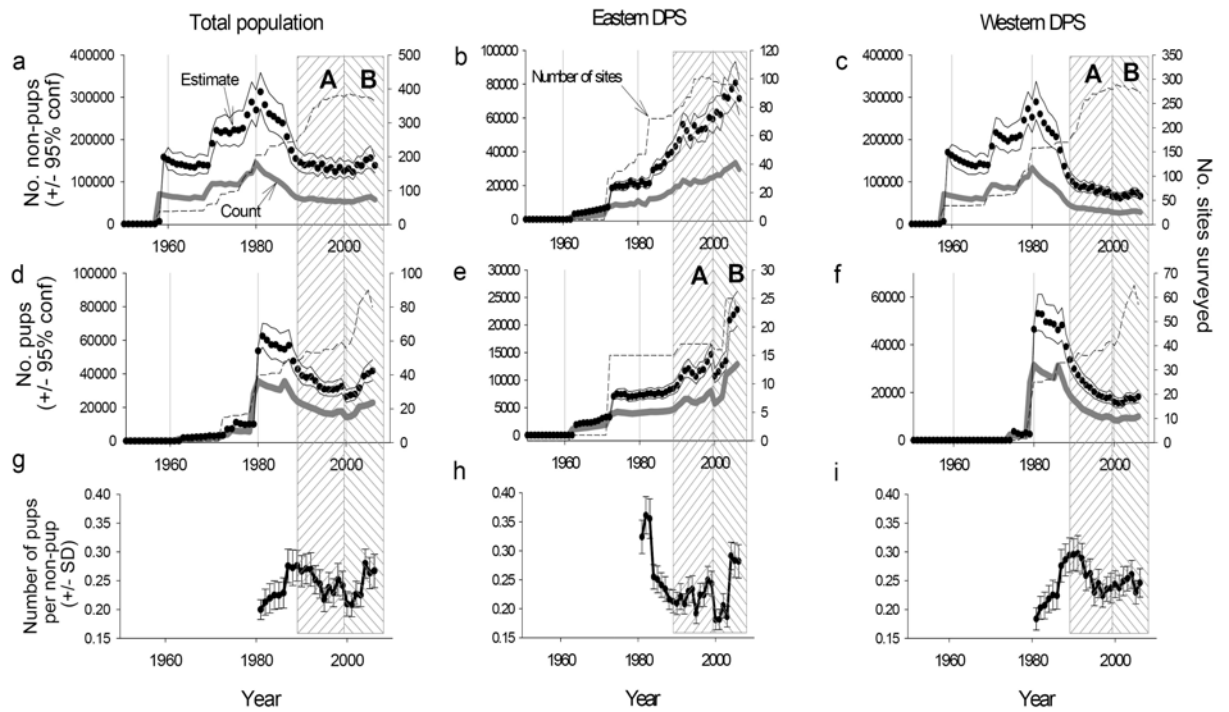


Figure 4

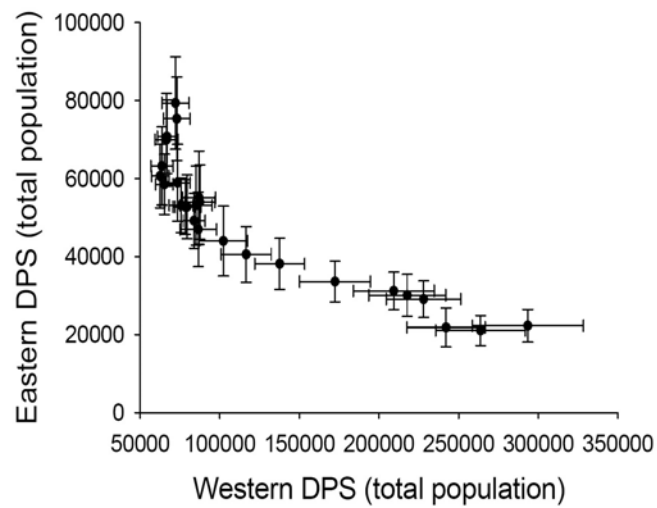


Figure 5

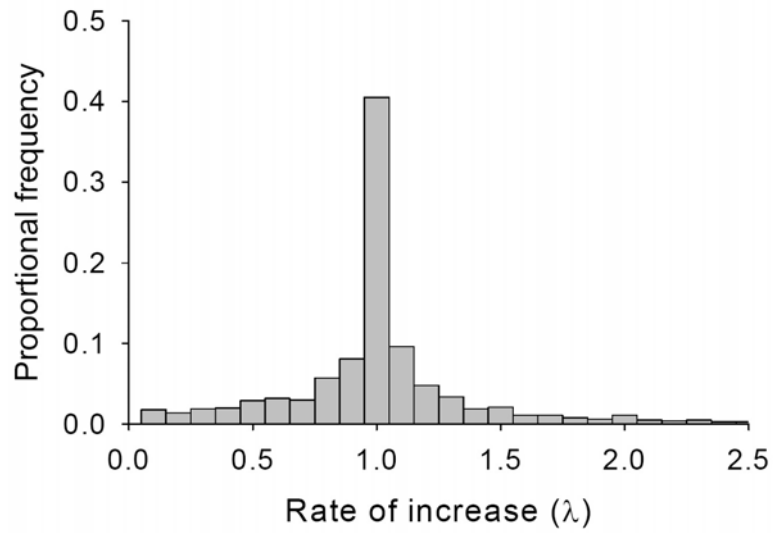


Figure 6

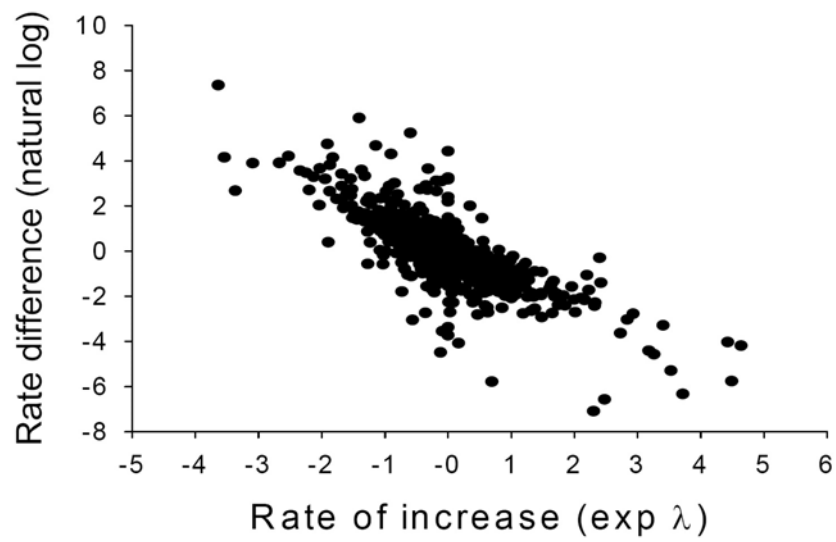


Figure 7

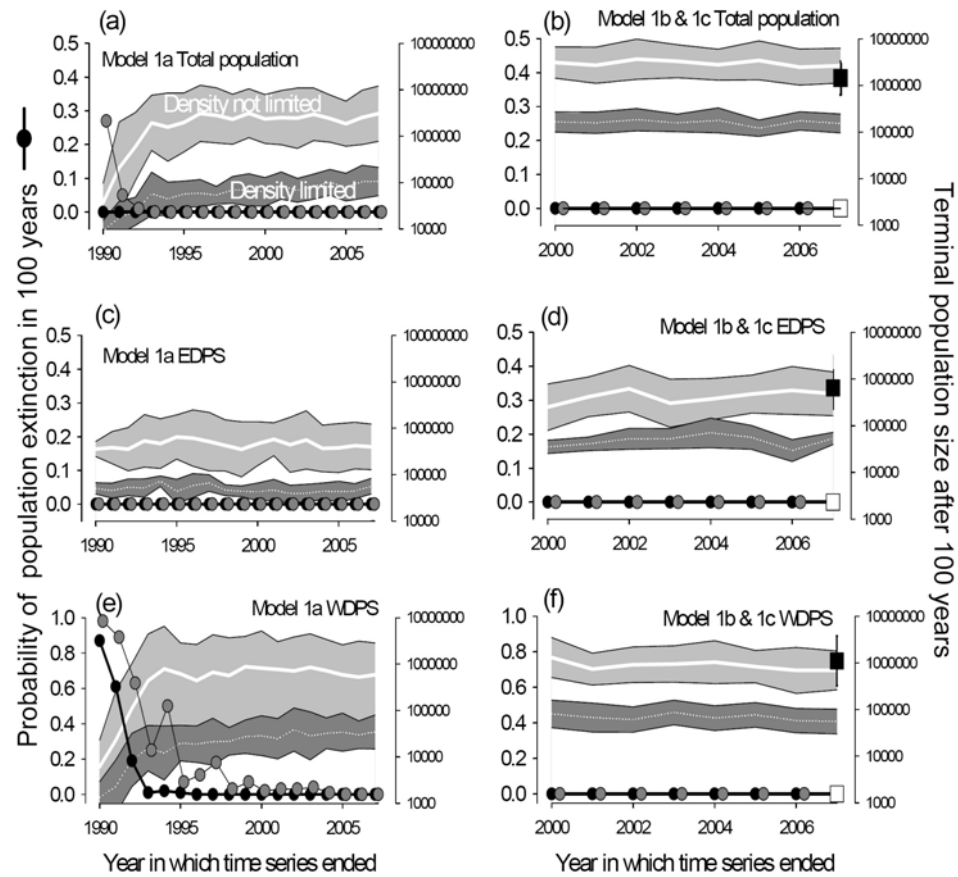
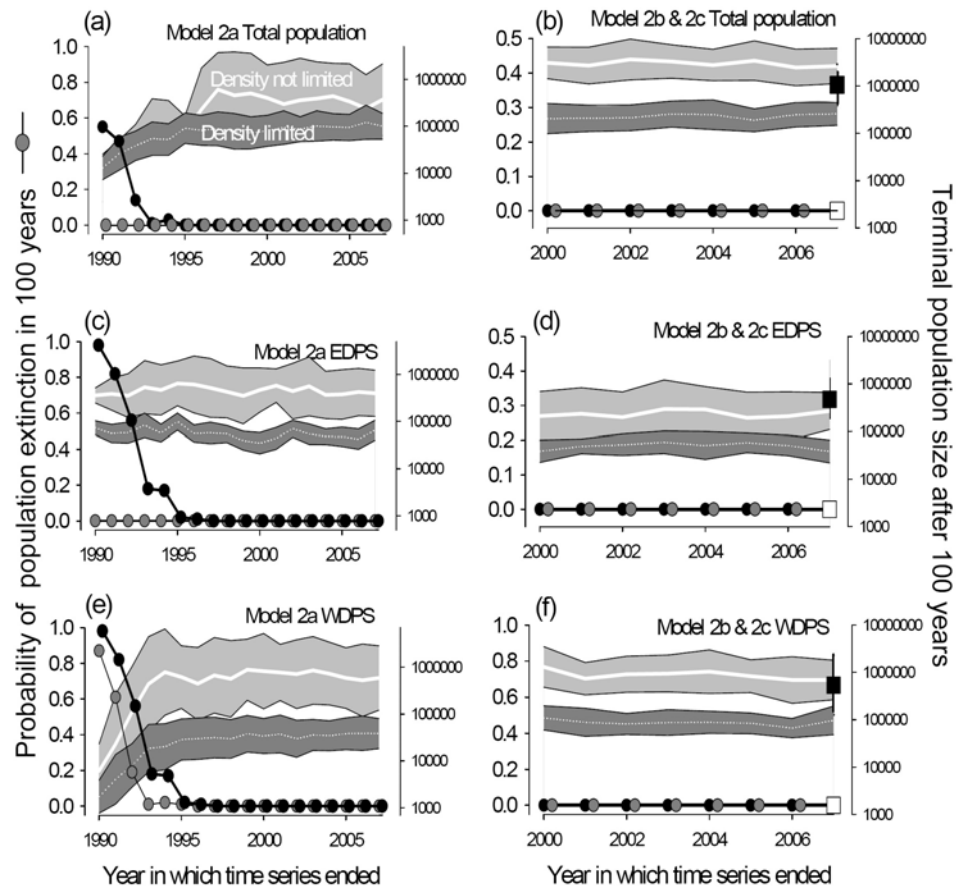


Figure 8



## **Assessing the effectiveness of conservation measures for the Steller sea lion (*Eumetopias jubatus*) – Supplementary material**

**I.L. Boyd**

Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews, St Andrews KY16 8LB, UK. [ilb@st-andrews.ac.uk](mailto:ilb@st-andrews.ac.uk)

### **Source of error within historical counts of Steller sea lions**

1. *Incomplete coverage* of the region by photography or by direct observation which excludes portions of the site from a survey. This is likely to be particularly important in terrain such as complex coastal topography. Although colonies of Steller sea lions tend to be reasonably well defined geographically, pre-conceptions of the extent of colonies could lead to some parts of colonies being missed, or some sites being missed altogether, if they have only recently formed and are not included within the survey route. This problem applies to all types of assessment methods. Aerial photographic methods have the disadvantage that it is not possible to link the counting of animals in real time to the process used to ensure there is complete coverage of the site. Oblique aerial photographs are also not geo-referenced with respect to one-another meaning that it is not possible to check post-hoc if some regions within the site have been missed. Other direct counting methods where there is a direct linkage between the counting process and the coverage process may be less susceptible to missing regions within sites but could also be constrained by an incapacity to observe some regions within sites by the characteristics of the topography. It is very difficult to know the extent to which this type of error can affect counts. Consequently, a broad range has been allocated in Table 1 and these have been varied according to an assessment of how easy it would be to ensure there was systematic coverage from sites using each method.

2. *The capacity to observe the animals* if they are present and if they have been included within the survey region. In the case of aerial photographs poor photograph quality is a substantial challenge in this respect. Large format photography and motion compensation introduced in later parts of the time series are particularly important to reduce this effect. The recent move to digital photography (Fritz et al. 2008) may reduce the likelihood of poor photograph quality but, in general, digital images do not currently yet have the definition of large format film. However, photograph quality will be a particular problem for the observation of animals when using hand-held small-format (35 mm) cameras with an oblique aspect. Different constraints on observation will apply to direct observation where there will be a trade-off between exhaustive searching for animals within a site and the time taken to



search. Three main sub-processes will operate as a result of observability issues:

(i) *Misclassification*. In general, misclassification is likely to be comparatively low because of the coarse classification process (between “pups” and “non-pups”) and because of the spatial separation of pups on rookeries from juveniles which are likely to be the main source of confusion with pups.

(ii) *False positive*. This occurs when an individual is recorded as being observed when it is not a Steller sea lion. False positives may occur because of confusion between animals and topographic features or other species, such as California sea lions where ranges overlap, because of poor photographic quality or poor viewing conditions involving low light levels or an unfavourable aspect. It may also occur when there is double-counting of animals because of undetected overlap between photographs.

(iii) *False negative*. This occurs when an individual is not observed when it is actually present. In general, this is likely to have a higher prevalence than false positives. It can occur because of poor photographic quality resulting in loss of definition, poor viewing aspect, topographic features that shield animals from view, larger animals obscuring the view of smaller animals.

It is difficult to fully quantify misclassification, false positive and false negative counts but they are likely to be highly site-, method- and time dependent. They will also vary between pups and non-pups. The ranges applied in Table 1 are best guesses at the level of this type of error.

3. *Variance due to counters*. Even in closely controlled conditions, different individual observers, whether counting photographs or making direct counts, will observe different numbers of animals. Using data from Fritz et al. 2008, it is possible to show that the effects of the counter on the numbers counted involved an average error of 0.9% for digital photography and 0.2% for large format film photography. Binomial tests show that the difference between digital and film photography when carried out using exactly the same methods in every other way is significant ( $p < 0.001$ ). We would expect digital and large-format photography to have minimised counter error and indeed the levels of error in this case are small. However, this also shows that, there is method-dependency in the counting error even for these highly refined survey methods. We would expect counter error to be considerably greater for the other counting methods and the error allocated to this source has been scaled in relation to the ease of counting for counters (Table 1)

4. *Time/date dependency*. The surveys are timed to take place from the middle of June to the middle of July, and between the 0900 and 1800h local time, when there are likely to be the maximum numbers of sea lions ashore (Pitcher et al. 2001; Sease and Gudmundson 2002). The coefficient of variation

reported among sites counted 4-11 days apart was from 0.029 to 0.044 (representing a range of count variation of 5%-10%) depending upon the region considered (Sease and Gudmundson 2002). Although these CVs capture counter variation as well as variation caused by differences in the timing of surveys, they represent an 11-day segment of the overall survey period used for Steller sea lions of about 30 days. Pitcher et al. (2001) reported that 90% of births at any specific site occur within 25 days and that the mean birth date ranges from 4 June to 21 June. Clinal variation in birth date, as well as inter-annual variation within sites of as much as 10 days, also occur. Consequently, additional variance is likely because of the timing of surveys in relation to the ambient seasonality at each site. For example, applying the results from Pitcher et al. (2001) to the surveys of Sease and Gudmundson (2002), in which pups were surveyed from 22 June to 9 July, shows that at some sites ~75% of pups could have been born at the time of surveying. Although the CVs provided by Sease and Gudmundson (2002) suggest this was not the case, these only applied to counts of non-pups. Olesiuk (2008) considered that 98-100% of birth would have occurred at the time of the surveys conducted in British Columbia but no variance has been attached to this estimate and it is unclear how timing affected the number of non-pups counted; CVs of 5-10% were present within counts irrespective of any effects of bias caused by timing or date dependency.

Chumbley et al. (1997) also quote CVs of overview counts (count made of beaches from vantage points) that were 8-12% during the period of peak abundance of non-pups at two beaches on Marmot Island but were 15-40% at a third beach. Chumbley et al. (1997) also compared directly between counts made from vantage points with those made using oblique aerial survey and, when counts were made simultaneously estimated that vantage point counts were 22% lower than the aerial survey counts suggesting a bias in vantage point counts.

5. *Terrain dependency.* Since the CVs supplied by Sease and Gudmundson (2002) were derived from repeated counts of the same site, they say nothing about variation caused by different forms of terrain. Steller sea lions breed on many different forms of coast and some features of the terrain may lead to reductions in the number of sea lions counted. This effect needs to be distinguished in some cases from the effects resulting from low observability, for example because animals are obscured by others, because they are obscured by rocks or overhanging cliffs, or because of poor image quality. Although under some types of observation methods in Table 1, terrain effects have been grouped with the observability of animals, in others, it was considered necessary to separate these for photographic methods.

A further class of estimation method has also been used on Steller sea lions. This is largely undefined but will likely vary from counts made visually and with some level of systematic coverage to counts made using a visual reference of the total number of animals observed without carrying out a count of individuals (as could happen when flying past a site). A high level of variation has been allocated to these methods with an assumption that there was generally an under-estimation of the total number, mainly because of the combined effects of all these other sources of error already listed.

### **Modelling spatial and temporal correlation between sites**

The objective of the analysis to determine spatial correlation was to define the extent to which the future trends at each site could be modelled independently or as part of a cluster of sites. The complexities of the data variance-covariance matrix, due to missing values, did not permit the use of standard statistical procedures. As a result a multiple regression approach was used to cluster sites based upon congruence between the value of the historical population trajectories at each site. This approach made use of the relationship:

$$\mathbf{b} = \mathbf{S}^{-1} \mathbf{y} \quad 5.$$

Where  $\mathbf{S}$  was the  $i \times i$  symmetrical matrix containing the sum of squares-sum of products matrix for a set of sites defined as the independent variables in a multiple regression. The vector  $\mathbf{y}$  of length  $i$  was the sum of products for the site designated as the dependent variable. In each case, the years included in the data set were those in which the sites in  $\mathbf{S}$  contained no missing values during the years in  $\mathbf{y}$  when there were also no missing values. The number of sites included as independent variables to derive  $\mathbf{S}$  was defined to be less than the number of years represented by the data. The minimum number of years of data required for a site to be included was 6. In order to maximise the power of this analysis log-linear interpolation was also used to estimate the number of animals present in years between counts and to reduce the effects of missing data. This allowed the estimation of  $\mathbf{b}$  which was a vector containing the partial regression coefficients for each site designated as an independent variable.

A systematic analysis was conducted using each site as the independent variable and all permutations of different sets of all other 431 sites. The goodness of fit of each model was assessed using the AIC and the best-fitting model was chosen for each site. Note that the AIC was only used to compare among models using the same dependent variable this avoiding the effects of over-parameterisation due to serial correlation. This had the effect of selecting a small set of sites that were statistically most closely related to each site, in

terms of their population dynamics, with either a positive or a negative relationship.

The sets of sites used as independent variables were then inspected to determine which sites occurred with greatest frequency as independent variables in the most favoured models. The correlation between these sites and other sites was preserved in the population projections of the dependent sites. Those sites for which there was no significant relationship with a trend site were treated as behaving independently of all other sites in projections.

Table 1. The total number of sites and individuals counted using different methods. The most probable sources of error or bias in the estimates are listed against each method and the effect of the errors, in terms of a positive or negative bias or of random error is given. In most cases, the magnitudes of the effects have not been measured formally. Therefore, effect was translated in to a magnitude expressed as a coefficient of variation of 0.5%, 1%, 2%, 5%, 10%, 15%, 20% or 50% change in the estimate.

Count method	Sites		Individuals		Source of error or bias	Effect	Magnitude	
	Pup counts	Non-pup counts	Pup counts	Non-pup counts				
Oblique aerial photography	0	4364	0	1,286,214	Photographic coverage of area because of no formal geo-referencing	double-counting	+1%	+5%
						animals missed	-20%	-60%
					Poor photograph quality	mis-classification	±5%	±15%
						false +ve	+1%	+5%
						false -ve	-5%	-10%
					Counter	+ve or -ve	±2%	±5%
					Date/time dependence	-ve	-2%	-20%
					Terrain dependence	-ve	-2%	-10%
Vertical medium-format aerial photography	362	1429	30,746	204,471	Incomplete photographic coverage of area	animals missed	-20%	-60%
					Poor photograph quality	mis-classification	±2%	±5%
						false +ve	+1%	+2%

						false -ve	-2%	-5%
					Counter	+ve or -ve	±0.5%	±2%
					Date/time dependence	-ve	-1%	-20%
					Terrain dependence	-ve	-2%	-10%
Vertical digital photography	2	0	22	0	Incomplete photographic coverage of area	animals missed	-20%	-60%
					Poor photograph quality	mis-classification	±2%	±10%
						false +ve	+1%	+2%
						false -ve	-2%	-5%
					Counter	+ve or -ve	±0.5%	±2%
					Date/time dependence	-ve	-2%	-20%
					Terrain dependence	-ve	-2%	-10%
Aerial visual counts	0	203	0	55,079	Incomplete coverage of site	animals missed	-20%	-70%
					Double-counting	+ve	+2%	+10%
					Observability of animals (includes terrain effects)	mis-classification	±5%	±20%
						false +ve	+5%	+10%
						false -ve	-5%	-15%
					Counter	+ve or -ve	±5%	±15%
					Date/time dependence	-ve	-2%	-20%
Beach/ground counts	232	128	101,466	12,201	Incomplete coverage of site	animals missed	-20%	-70%

					Double-counting	+ve	+2%	+10%
					Observability of animals (includes terrain effects)	mis-classification	±0.5%	±0.5%
					Counter	+ve or -ve	±2%	±5%
					Date/time dependence	-ve	-2%	-20%
Viewpoint counts	19	18	5,867	3,270	Incomplete coverage of site	animals missed	-20%	-70%
					Double-counting	+ve	+2%	+5%
					Observability of animals (includes terrain effects)	mis-classification	±0.5%	±0.5%
						false +ve	+1%	+2%
						false -ve	-5%	-15%
					Counter	+ve or -ve	±1%	±2%
					Date/time dependence	-ve	-2%	-20%
Counts from skiff	33	135	2,934	57,146	Incomplete coverage of site	animals missed	-30%	-80%
					Double-counting	+ve	+2%	+10%
					Observability of animals (includes terrain effects)	mis-classification	±2%	±5%
						false +ve	+1%	+2%
						false -ve	-10%	-20%
					Counter	+ve or -ve	±2%	±5%
					Date/time dependence	-ve	-2%	-20%
Counts from vessel	8	0	194	0	Incomplete coverage of site	animals missed	-20%	-80%

offshore					Double-counting	+ve	+2%	+10%
					Observability of animals (includes terrain effects)	mis-classification	±2%	±5%
						false +ve	+1%	+2%
						false -ve	-10%	-20%
					Counter	+ve or -ve	±2%	±5%
					Date/time dependence	-ve	-2%	-20%
Undefined/other	0	13	158,876	342,927	Non-specific effects; counting methods generally non-systematic	+ve or -ve	-50%	+20%